

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-347692

(43)Date of publication of application : 15.12.2000

(51)Int.Cl.

G10L 17/00  
G06T 1/00  
// A61B 3/113

(21)Application number : 11-159882

(71)Applicant : SANYO ELECTRIC CO LTD

(22)Date of filing : 07.06.1999

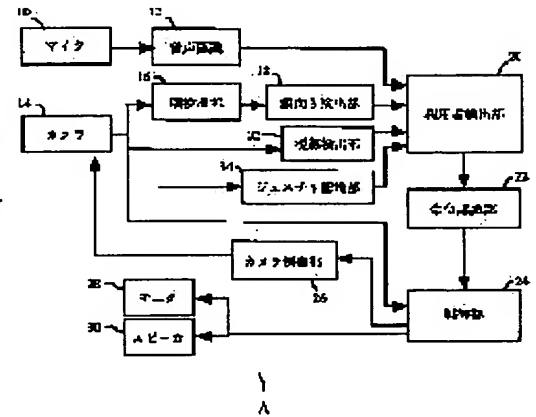
(72)Inventor : HONGO HITOSHI

(54) PERSON DETECTING METHOD, PERSON DETECTING DEVICE, AND CONTROL SYSTEM USING IT

(57)Abstract:

**PROBLEM TO BE SOLVED:** To specify the user of a device within a plurality of persons by detecting a face image from an image photographed by a camera to estimate the direction of one's eye or a face, and detecting a speaking person as the user.

**SOLUTION:** When voice is input to a microphone 10, voice recognition is performed, and a face territory is detected from the image input from a camera 14. Against the detected face territory, the direction of the face is detected, the direction or the object observed by the person is estimated, and whether the person observes in the specified direction or object or not is judged. In the case of YES, the person is judged to be a user, the content indicated or ordered by the detected user is recognized, and when it is indication by voice, the result of voice recognition is converted into a command fit for the device. When the user indicates by gesture, the result recognizing the gesture is converted into a command. Further, for judging the user, not only the direction of face but the eye direction is detected, and it is enough to judge whether the eye direction gazes at the object or not.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2000-347692  
(P2000-347692A)

(43) 公開日 平成12年12月15日 (2000. 12. 15)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーム (参考)
G 1 0 L 17/00		G 1 0 L 3/00	5 4 5 F 5 B 0 5 7
G 0 6 T 1/00		G 0 6 F 15/62	3 8 0 5 D 0 1 5
// A 6 1 B 3/113		A 6 1 B 3/10	B 9 A 0 0 1

審査請求 未請求 請求項の数13 O L (全 6 頁)

(21) 出願番号 特願平11-159882

(22) 出願日 平成11年6月7日 (1999. 6. 7)

(71) 出願人 000001889

三洋電機株式会社

大阪府守口市京阪本通2丁目5番5号

(72) 発明者 本郷 仁志

大阪府守口市京阪本通2丁目5番5号 三

洋電機株式会社内

(74) 代理人 100111383

弁理士 芝野 正雅

Fターム (参考) 5B057 DA06 DA07 DB03 DC02 DC08

5D015 AA03 CC01 CC12 HH00 LL07

9A001 HH17 HH23 HH30 HH34 JJ23

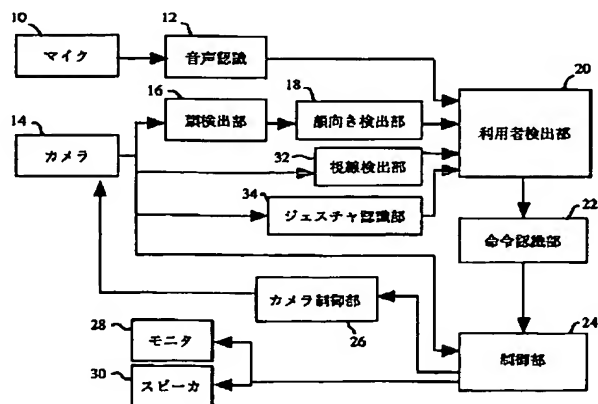
KK56

(54) 【発明の名称】 人物検出方法、人物検出装置及びそれを用いた制御システム

(57) 【要約】

【課題】 複数の人物から装置の利用者を適切に検出することで、複数の人物が居る環境下でも利用できるインタフェースを備えた装置を提供すること。

【解決手段】 複数の人物から利用者を検出する方法として、カメラで撮影した映像から顔画像を検出する手段と、検出した顔画像の視線方向または顔の向きを推定し、その推定結果からある所定の方に視線または顔を向けいるかを判断する手段と、ある所定の方に視線または顔を向け、かつ発話している人物を利用者として検出する手段と、を備える。



【特許請求の範囲】

【請求項1】 利用者の音声情報、操作情報、画像情報のうち、少なくとも一つの情報に基づき複数の人物から利用者を検出することを特徴とする人物検出方法。

【請求項2】 請求項1に於いて、前記音声情報が、音声、または音声認識により得られた情報であり、発声、口調、テンポ、声紋、言語のいずれかであることを特徴とする人物検出方法。

【請求項3】 請求項1に於いて、前記操作情報が、キーボード、マウス、リモコン、ペン、タッチパネルのいずれかによる操作で入力される情報であることを特徴とする人物検出方法。

【請求項4】 請求項1に於いて、前記画像情報が、映像、または画像認識により得られた情報であり、顔、顔の向き、視線方向、瞬き、うなずき、唇の動き、ポーズ、ジェスチャのいずれかであることを特徴とする人物検出方法。

【請求項5】 請求項1に於いて、前記画像情報が映像であり、映像から顔画像を検出し、検出した顔画像の顔の向き、または視線方向を推定し、その推定結果からある所定の方向に顔または視線を向け、且つ発話している人物を利用者として検出することを特徴とする人物検出方法。

【請求項6】 請求項5に於いて、前記画像情報として、更に唇の動きを検出して、ある所定の方向に顔または視線を向けながら発話し、且つ口が動いている人物を利用者として検出することを特徴とする人物検出方法。

【請求項7】 請求項5に於いて、前記画像情報として、更に手の動きを検出し、ある所定の方向に顔または視線を向けながらジェスチャしている人物を利用者として検出することを特徴とする人物検出方法。

【請求項8】 利用者の音声情報、操作情報、画像情報のうち、少なくとも一つの情報に基づき複数の人物から利用者を検出する手段を備えたことを特徴とする人物検出装置。

【請求項9】 カメラで撮影した映像から顔画像を検出する手段と、検出した顔画像から視線方向または顔の向きを推定する手段と、前記顔画像がある所定の方向に視線または顔を向けていることを判断する手段と、前記顔画像の人物の音声情報を検出する手段と、を備えたことを特徴とする人物検出装置。

【請求項10】 カメラで撮影した映像から顔画像を検出する手段と、検出した顔画像から視線方向または顔の向きを推定する手段と、前記顔画像がある所定の方向に視線または顔を向けていることを判断する手段と、前記顔画像の人物の音声情報を検出する手段と、前記顔画像から唇の動きを検出する手段と、を備えたことを特徴とする人物検出装置。

【請求項11】 カメラで撮影した映像から顔画像を検出する手段と、検出した顔画像から視線方向または顔の

向きを推定する手段と、前記顔画像がある所定の方向に視線または顔を向けていることを判断する手段と、前記映像から前記顔画像の人物によるジェスチャを検出する手段と、を備えたことを特徴とする人物検出装置。

【請求項12】 利用者の音声情報、操作情報、画像情報のうち、少なくとも一つの情報に基づき複数の人物から利用者を検出する手段と、この検出手段の検出結果に基づき前記利用者をカメラを用いて顔を撮影する手段と、前記利用者の撮影している顔をモニタに出力する手段を備えた制御システム。

【請求項13】 利用者の音声情報、操作情報、画像情報のうち、少なくとも一つの情報に基づき複数人の中から利用者を検出する手段と、この検出手段の検出結果に基づき制御信号を出力する制御手段と、前記制御手段を一人の利用者からの命令を受けるモードまたは、複数人の命令を並列で処理して受けるモードに切り替える手段とを備えた制御システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、複数の人物が居る環境下で、利用者を検出し、その人物と対話する装置に適用して最適な人物検出方法、人物検出装置、及び制御システムに関する。

【0002】

【従来の技術】従来から装置の利用者、例えば、発言者または操作者を検出する方法として、主に音声を用いられてきた。特開平5-122689号公報では、複数のマイクから入力される音声を検出し、音声入力があったマイクの方向にカメラを向ける方法が開示されている。また、特開平6-351015号公報、特開平7-140527号公報では、指向性マイク、あるいはマイクアレーを用いて音から発話者の方向を検出し、その方向にカメラを向ける発明を開示している。

【0003】

【発明が解決しようとする課題】しかし、上記のように音による方法では、複数の人が同時に話をしている場合、その中から装置の利用者を適切に検出することは困難である。また、利用者の近傍に壁や物があった場合、声の反響によって利用者の検出を誤る恐れがあった。また、ジェスチャなど音を発しない方法では、これらの装置を利用することができなかった。

【0004】そこで、本発明では、装置に対して利用している人物を適切に検出することで、複数の人物が居る環境下でも利用できるインタフェースを備えた装置を提供することを目的とするものである。

【0005】

【課題を解決するための手段】本発明は、上記問題を解決するために創作されたものであって、第1には、カメラで撮影した映像から顔画像を検出する手段と、前記顔画像から視線方向または顔の向きを推定し、その推定結

果からある所定の方向に視線を向けている、または顔を向けているかを判断する手段と、前記顔画像の人物が発話しているかを検出する手段と、を有することを特徴とする。

【0006】この第1の構成の人物検出方法および人物検出装置に於いては、カメラで撮影した映像から顔画像を検出し、検出した顔画像から視線または顔の向きを推定し、その推定結果から操作する対象物に視線または顔を向け、且つその人物が発話している場合、その人物を利用者として検出する。よって、操作する対象物を見ながら発話しているかを判断することで複数の人物の中から利用者を特定することができる。

【0007】また、第2には、カメラで撮影した映像から顔画像を検出する手段と、前記顔画像から視線方向の検出手段と、前記顔画像から視線方向または顔の向きを推定し、その推定結果がある所定の方向に視線または顔を向けているかを判断する手段と、前記顔画像の人物が発話しているかを検出する手段と、前記顔画像の人物の唇が動いているかを検出する手段と、を有することを特徴とする。

【0008】この第2の構成の人物検出方法および人物検出装置に於いては、カメラで撮影した映像から顔画像を検出し、検出した顔画像から視線または顔の向きを推定し、その推定結果から操作する対象物に視線または顔を向けて、且つその人物が発話し、且つ唇が動いている場合、その人物を利用者として検出する。よって、装置の利用者と異なる人物が発話してもその発話が利用者の命令かを見極めることができ、複数の人物の中から利用者を的確に特定することができる。

【0009】また、第3には、カメラで撮影した映像から顔画像を検出する手段と、前記顔画像から視線方向の検出手段と、前記顔画像から視線方向または顔の向きを推定し、その推定結果がある所定の方向に視線または顔を向けているかを判断する手段と、前記顔画像の人物によるジェスチャを検出する手段と、を有することを特徴とする。

【0010】この第3の構成の人物検出方法及び人物検出装置に於いては、カメラで撮影した映像から顔画像を検出し、検出した顔画像から視線方向または顔の向きを推定し、その推定結果から操作する対象物に視線または顔を向けているかを判断し、且つ装置の操作主導権を得るために、または装置に命令を伝達するために、例えば手を挙げるなどのジェスチャを検出することで利用者と判定を行う。よって、操作する対象物を見ている人物で且つ装置の操作主導権を要求するジェスチャ、または装置への命令を表すジェスチャから利用者を判断することで複数の人物の中から利用者を特定することができる。

【0011】また、第4には、利用者の音声情報、操作情報、画像情報のうち、少なくとも一つの情報に基づき複数の人物から利用者を検出する手段と、この検出手段

の検出結果に基づき前記利用者をカメラを用いて顔を撮影する手段と、前記利用者の撮影している顔をモニタに出力する手段と、を有することを特徴とする。

【0012】この第4の構成の制御システムに於いては、利用者の音声情報、操作情報、画像情報のうち、少なくとも一つの情報に基づき複数の人物から利用者を検出し、利用者にカメラを向け、カメラでとらえた利用者の顔画像をモニタに表示することで誰が利用者であるかをその場にいる人物に知らせる。よって、制御システムが利用者の顔画像を表示することで、複数の人物の中から誰が利用者として選択されているかを知ることができる。

【0013】また、第5には、利用者の音声情報、操作情報、画像情報のうち、少なくとも一つの情報に基づき複数人の中から利用者を検出する手段と、この検出手段の検出結果に基づき制御信号を出力する制御手段と、前記制御手段を一人の利用者からの命令を受けるモードまたは、複数人の命令を並列で処理して受けるモードに切り替える手段とを有することを特徴とする。

【0014】この第5の構成の制御システムに於いて、複数人物の中から利用者だけの命令情報を処理するモード、または複数の人物の命令を並列に処理するモードを設け、これらのモードを切り替えられるようにする。よって、前記制御システムの操作主導権を他人に取られないようにすることができる。また、複数の人物が同時に指示・命令を出すことができる。

【0015】

【発明の実施の形態】本発明の実施の形態としての実施例を図面を利用して説明する。

【0016】本発明に基づく人物検出装置Aは、図1に示されるように、マイク10と、音声認識部12と、カメラ14と、顔検出部16と、顔向き検出部18と、利用者検出部20と、命令認識部22と、制御部24と、カメラ制御部26と、モニタ28と、スピーカ30と、視線検出部32と、ジェスチャ認識部34と、を有している。

【0017】ここで、音声入力手段としての上記マイク10は、入力された音声を電気信号としての音声信号に変換するものである。また、音声認識部12は、上記マイク10から入力された音声信号の波長等を分析することにより入力された音声を認識して、所定の音声データ（単語）を登録しておくことで認識精度を向上させることができる。

【0018】また、上記撮影手段としての上記カメラ14は、利用者を撮影するものである。このカメラ14は、例えば、CCDカメラ等により構成される。また、上記顔検出部16は、上記カメラ14により得られた画像データから顔領域を検出する。上記顔向き検出部18は、上記顔検出部16により検出した顔領域から顔の向

きを算出する。

【0019】また、上記視線検出部32は、利用者の視線方向を算出する。また、上記ジェスチャ認識部34は手振り、身振りなどの動作や手によるサイン、ポーズを認識する。

【0020】また、上記利用者検出部20は、上記音声認識部12、上記顔向き検出部18、上記視線検出部32、上記ジェスチャ認識部34の少なくとも1つから検出情報が送られ、命令を発した人物、あるいは所定の方向を見ながら命令を発している人物、あるいは所定の方向を見ながら命令を発し、且つ口が動いている人物、あるいは所定の方向を見ながら所定のジェスチャまたはポーズを示している人物を利用者と判断する。

【0021】上記命令認識部22は、上記利用者検出部20で利用者と判断した人物の音声による命令、またはジェスチャによる命令を認識し、上記制御部24に対して所定の命令を発信する。上記制御部24は、上記命令認識部22からの発信された命令に従い所定の実行処理を行うものである。上記モニタ28、上記スピーカは、装置が処理した結果を出力するためのものである。

【0022】また、上記カメラ制御部26は、上記利用者検出部20で検出した利用者の位置からカメラ方向を算出し利用者の方向にカメラを向けるための制御を行う。

【0023】また、上記モニタ28は、上記カメラ16と上記顔検出部16で得られた情報から、利用者の顔画像をモニタに出力する、あるいは子画面で表示することで現在の利用者を知らせる。

【0024】なお、上記人物検出装置Aのうち、マイク10と、音声認識部12と、カメラ14と、顔検出部16と、顔向き検出部18と、利用者検出部20、命令認識部22が上記人物検出装置として機能する。本実施例では、利用者を検出するためにマイク10と、音声認識部12と、カメラ14と、顔検出部16と、顔向き検出部18を構成するものと説明したが、これには限られず、人物を検出する手段であれば他の構成としてもよい。

【0025】上記構成に基づく人物検出装置Aの動作について、図2を利用して説明する。なお、この場合には、図3に示すように、上記人物検出装置Aがテレビジョン受信装置Pに搭載されているものとして説明する。

【0026】まず、マイク10に音声入力があるかを判断する(S10)。そして、音声入力がある場合には、ステップ11に移行し、音声認識を行う。

【0027】ステップ12では、カメラ14から入力された画像から顔領域を検出する。顔領域を検出する方法としては、例えば、本出願人が特願平10-291543号で提案した方法を用いることができる。ステップ12に於いて、顔領域が検出されれば、ステップ13に進み、顔領域が検出されなければ、ステップ10に戻る。

ステップ13に於いて、ステップ12で検出した顔領域に対して顔の向きを検出する。顔の向きを検出する方法としては、例えば、本出願人が特願平10-269600号で提案した方法を用いることができる。

【0028】ステップ14に於いては、ステップ12で検出した人物がどの方向を見ているか判断する。このとき、前記人物が観察している対象物を知るためには、前記人物の位置を知る必要がある。検出した人物の位置を求める方法として、カメラの焦点距離を用いたり、予め人物の大きさとカメラからの距離の関係を対応して記録しておくことで、例えば顔の大きさからその人物の位置を求めることができる。あるいは、カメラ14を複数台のカメラにすることで、ステレオマッチングにより距離を求めることができる。あるいは、ここでは図示していないが距離センサー、赤外線センサなどを用いて距離、位置を求めることもできる。ステップ14では、検出した人物が観察している方向、または物を推定し、所定の方向、または対象物を見ているかを判断する。見ていると判断したらその人物を利用者と判断し、ステップ15に進み、見ていないと判断したらステップ10に戻る。

【0029】なお、上記の例では、利用者の判定に顔の向きを利用したが、人物の視線方向を検出し、視線方向が対象物の方を見ているかを判断してもよい。

【0030】ステップ15では、ステップ14で検出した利用者が指示・命令した内容を認識し、音声による指示ならば、音声認識の結果をその装置に適合した命令(コマンド)に変換する。利用者がジェスチャで指示したならば、ジェスチャを認識した結果を命令に変換する。身振り、手振りなどのジェスチャを認識したり、ジェスチャを登録したりする方法としては、例えば、“動作者適応のためのオンライン教示可能なジェスチャ動画像のスポテティング認識システム”、(電子情報通信学会論文誌、Vol. J81-D-II、No. 8、pp1822-1830)などに示された方法を利用して実現することができる。なお、ジェスチャと命令の対応付けは、予め登録または変更することができる。

【0031】ステップ16では、変換した命令を装置に送信し、ステップ17ではその命令を実行する。

【0032】なお、上記の例に於いて、音声入力と顔向き検出のみで音声による命令を実行すると、装置に顔を向けている人物と音声を発した人物とが別の場合でも動作してしまう惧れがある。そこで、カメラ10から得た画像から検出された利用者の口の動きがあるかどうかについても判定することが好ましい。すなわち、この場合には、図2のフローチャートにおけるステップS14に於いて、顔が所定の方向に向いているかどうかについて、更に、口の動きがあるかどうかについてが判定される。つまり、音声入力があり、顔向きが検出されても、口の動きが検出されなければ、利用者と判断しない。

【0033】上記のようにすることで装置に顔を向けて

いる人物とは別の人物が音声を発した場合でも、利用者と無関係な音声を排除して、誤って動作することを回避することができる。

【0034】なお、上記の例に於いて、人物検出装置Aを搭載する機器をテレビジョン受信装置Pとして説明したが、これには限られず、ビデオ、パソコン、エアコン、室内灯等の家電製品でもよく、他のあらゆる機器に搭載が可能である。

【0035】

【発明の効果】以上、本発明によれば、複数の人物から、利用者を的確に検出し、その利用者の指示・命令を認識することができる。

【0036】また、複数の人が同時に話している場合、また利用者の近傍に壁や物があることによる声の反響がある場合でもその中から利用者を適切に検出することができる。

【0037】また、ジェスチャなど音を発しない場合でも利用者を検出することができる。

【0038】また、利用者にカメラを向ける、あるいはカメラでとらえた利用者の顔画像を表示することで、誰が装置の操作指導権を持っているかを容易に把握させることができる。

【図面の簡単な説明】

【図1】本発明の実施例に基づく人物検出装置の構成を示すブロック図である。

【図2】本発明の実施例に基づく人物検出装置の動作を示すフローチャートである。

【図3】本発明の実施例に基づく人物検出装置の使用状況を示す説明図である。

【符号の説明】

A 人物検出装置

10 マイク

12 音声認識部

14 カメラ

16 顔検出部

18 顔向き検出部

20 利用者検出部

22 命令認識部

24 制御部

26 カメラ制御部

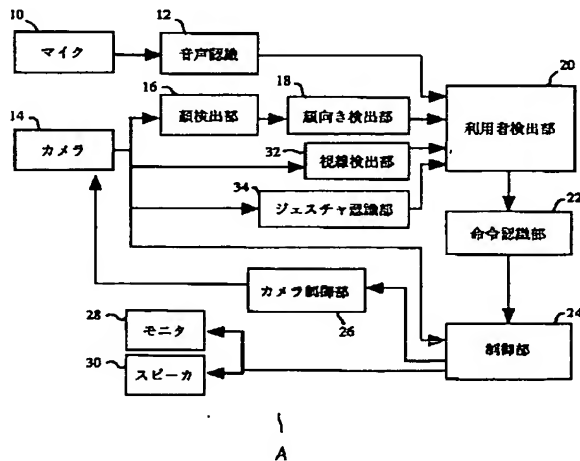
28 モニタ

30 スピーカ

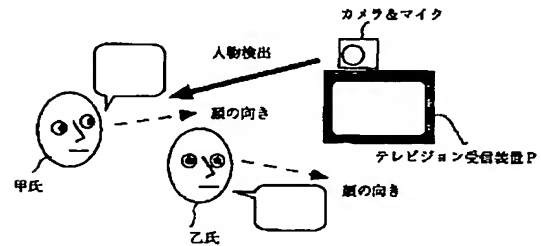
32 視線検出部

34 ジェスチャ認識部

【図1】



【図3】



【図2】

